

## **PENERAPAN TEORY ROUGH SET UNTUK MEMPREDIKSI TINGKAT KELULUSAN SISWA DALAM UJIAN NASIONAL PADA SMA NEGERI 5 KOTA BANDA ACEH**

Andika Prajana  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Ar Raniry Banda Aceh

**Abstract:** This research is used to set boundaries problem and object to be examined with the student's ability to identify and classify the data based on students' grades. After that the process of transformation of the data of the students in SMA Negeri 5 Kota Banda Aceh, Province of Nanggroe Aceh Darussalam consisting of average grades (UN), the value of Pre-UN and UN values. And analyze data on student scores using teory rough set, so finding a pattern or knowledge that can be used to predict the graduation rate of students in the next examination.

**Keywords:** *Data, Transformation, Rough Set, Prediction Graduation.*

### **PENDAHULUAN**

Salah satu kebijakan evaluasi pendidikan yang bersifat makro adalah dengan digulirkannya Ujian Nasional (UN). Ujian Nasional ini digulirkan dengan maksud untuk mengevaluasi hasil akhir belajar siswa dalam satu jenjang pendidikan tertentu dan sekaligus Ujian Nasional ini menjadi faktor evaluasi pendidikan yang menentukan apakah siswa lulus atau tidak lulus.

Peningkatan batas lulus 5.25 (tahun 2008) menjadi 5.50 (tahun 2013) oleh sebagian kalangan masyarakat/sekolah direspon dengan penuh kecemasan. Lazimnya setiap ada kenaikan standar kelulusan UN, selalu dianggap terlalu tinggi dan tidak memperhatikan kemampuan siswa yang berbeda-beda pada setiap daerah. Kaitannya dengan standar kelulusan UN sesungguhnya dapat dilakukan dengan beberapa cara/metode. Beberapa cara yang dapat digunakan yakni metode yang terfokus pada *judgement*, tes dan peserta didik (Heri Retnawati, dkk).

Selama ini penentuan batas lulus UN selalu dilakukan berdasarkan pertimbangan kebijakan (*judgement*) pemerintah

saja. Pola seperti ini, terbukti menimbulkan kecemasan karena masyarakat /sekolah tidak mengetahui mengapa sebuah nilai batas lulus (*standard setting/cut of score*) harus ditentukan.

*Teory Rough set* merupakan salah satu teknik yang dirasa cukup efisien untuk *Knowledge Discovery in Database* (KDD) Proses dan Data mining , (*Soft Computing Research Group*, UTM, 2001). Data mining merupakan serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data, (Iko Pramudiono, 2003)

### **Teori Rough Set**

*Rough Set* adalah sebuah teknik matematika yang dikembangkan oleh Pawlack pada tahun 1980. Teknik ini digunakan untuk menangani masalah *Uncertainty*. (*Missing data, Incompleted Data dan Inconsistency Data Imprecision dan Vagueness*) dalam aplikasi *Artificial Intelligence (AI)*.

Rough Set merupakan teknik yang efisien untuk *Knowledge Discovery* dalam Database (KDD) proses dan Data Mining. Secara umum, teori Rough Set telah digunakan dalam banyak aplikasi seperti *medicine, pharmacology, business, banking, engineering design, image processing dan decision analysis*. *Rough Set* merupakan teknik yang efisien untuk KDD proses dan *Data Mining*.

### **Atribut dalam Rough Set**

Pada sebuah database terdapat dua jenis atribut yaitu atribut kondisi dan atribut keputusan yang kadang disebut atribut kelas. Sedangkan pada *rough set* terdapat dua buah atribut yang digunakan yaitu :

#### **1. Atribut Kondisi (*conditional attribute*)**

Merupakan satu-satunya yang akan mewakili pengetahuan yang akan memelihara kedapatan dilihat antara objek-objek dengan menggunakan fungsi-fungsi kedapatan dilihat (Pawlak, 1991; Skowron dan Rauszer, 1992).

#### **2. Atribut Keputusan (*decision attribute*)**

Merupakan pengetahuan yang diungkapkan oleh salah satu atau

beberapa atribut dari atribut kondisi.  
Atribut keputusan berhubungan dengan konsep-konsep (kelas-kelas) yang tergambar dalam sebuah aturan. Biasanya, sebuah atribut keputusan yang tunggal sudah merupakan semua yang dibutuhkan.

1. Representasi Data dalam Rough Set

Terdapat dua bentuk representasi data dalam *rough set*, yaitu  
1. Information Systems (IS)

*Information system* adalah tabel yang terdiri dari baris yang merepresentasikan data (objek) dan kolom yang merepresentasikan atribut atau variabel dari data. *Information system* pada *data mining* dikenal dengan nama *dataset*.

*Information system* dapat direpresentasikan sebagai fungsi dimana:  
 $IS=\{U,A\}$ ..... (1)

$U=\{E_p, E_q, ..., E_m\}$ , adalah himpunan tidak kosong dari objek yang direpresentasikan dan  $A=\{a_1,a_2,..., a_n\}$ , adalah himpunan tidak kosong dari atribut atau variable.

Pada Tabel 1 memperlihatkan sebuah *information system* yang sederhana. Dalam *information system*, tiap-tiap baris merepresentasikan objek seperti yang terlihat bahwa *example* ( $U=S_1,S_2, .....S_{12}$ ), sedangkan kolom merepresentasikan atribut ( $A=$  Rata-rata Nilai Rapor(UN), Rata-rata Nilai Pra UN, Rata-rata Nilai UN). Di sini terlihat bahwa semua atributnya adalah *conditional attribute*.

Tabel 1 Information System

Example	RATA2 NILAI RAPOR (UN)	RATA2 NILAI PRA UN	RATA2 NILAI UN	KET
S <sub>1</sub>	6.95	5.78	7.81	LULUS
S <sub>2</sub>	8.29	7.28	7.60	LULUS
S <sub>3</sub>	6.46	8.51	7.85	LULUS
S <sub>4</sub>	6.24	4.64	6.75	LULUS
S <sub>5</sub>	6.24	5.29	7.25	LULUS
S <sub>6</sub>	6.03	5.34	6.66	LULUS
S <sub>7</sub>	6.06	4.08	6.74	LULUS

$S_8$	6.41	5.28	6.30	LULUS
$S_9$	7.15	5.88	7.01	LULUS
$S_{10}$	6.71	5.39	5.34	TIDAK LULUS
$S_{11}$	6.65	7.41	5.38	TIDAK LULUS
$S_{12}$	6.65	5.49	7.94	LULUS

Decision Systems (DS)

*Decision system* adalah *information system* dengan atribut tambahan yang dinamakan dengan *decision attribute*, dalam *data mining* dikenal dengan nama kelas atau target. Atribut ini merepresentasikan hasil dari klasifikasi yang diketahui.

*Decision system* merupakan fungsi yang mendeskripsikan *information system* yaitu, dimana :

$$DS = \{U, (A,C)\} \dots\dots\dots(2)$$

- U= objek
- A= atribut kondisi
- C= atribut keputusan

Dalambanyakaplikasi, sebuah *outcome* dari pengklasifikasian diketahui yang direpresentasikan dengan sebuah *decision attribute* ( C ). Maka *Information System* (IS) menjadi *Decision System* (DS)=(U,{A,C}). *Decision System* yang sederhana diperlihatkan pada Tabel 2.

Pada Tabel 2 memperlihatkan sebuah *Decision System* yang sederhana. Terdiri dari objek seperti  $S_1, S_2, \dots, S_{12}$  dan attribute kondisi seperti Rata-rata Nilai Rapor(UN), Rata-rata Nilai Pra UN, Rata-rata Nilai UN, serta atribut keputusan yaitu tingkat kelulusan.

Tabel 2 Decision System

Example	RATA2 NILAI RAPOR (UN)	RATA2 NILAI PRA UN	RATA2 NILAI UN	KET	TINGKAT KELULUSAN
$S_1$	6.95	5.78	7.81	LULUS	100%
$S_2$	8.29	7.28	7.60	LULUS	100%
$S_3$	6.46	8.51	7.85	LULUS	100%
$S_4$	6.24	4.64	6.75	LULUS	67%
$S_5$	6.24	5.29	7.25	LULUS	67%
$S_6$	6.03	5.34	6.66	LULUS	67%

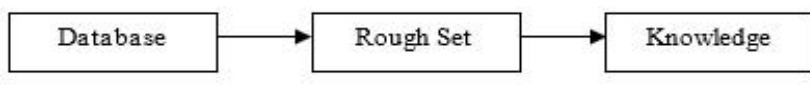
S <sub>7</sub>	6.06	4.08	6.74	LULUS	67%
S <sub>8</sub>	6.41	5.28	6.30	LULUS	67%
S <sub>9</sub>	7.15	5.88	7.01	LULUS	100%
S <sub>10</sub>	6.71	5.39	5.34	TIDAK LULUS	25%
S <sub>11</sub>	6.65	7.41	5.38	TIDAK LULUS	50%
S <sub>12</sub>	6.65	5.49	7.94	LULUS	67%

### Perancangan Umum Sistem Prediksi Tingkat Kelulusan Siswa dalam Ujian Nasional

Perancangan sistem dipakai untuk memberikan perencanaan dalam membangun sistem aplikasi data mining. Perancangan ini dimulai dari menentukan alur kerja sistem berikut data yang mengalir dalam sub sistem yang ada, dilanjutkan dengan merancang database, dan merancang proses-proses yang ada di setiap sub sistem.

Berdasarkan pada sub bab analisa dan sub bab evaluasi system yang sedang berjalan yang telah diuraikan di atas, maka secara sederhana proses penemuan *knowledge* yang akan digunakan dalam proses memprediksi tingkat kelulusan siswa dalam ujian nasional tersebut dapat dilihat pada Gambar 1.

**Gambar 1** Gambaran Sederhana Proses Penemuan *Knowledge* dengan *Rough Set*



Pada Gambar 1 terlihat bahwa proses penemuan *knowledge* dengan *rough set* terdiri atas tiga yaitu:

1. Database

Database merupakan kumpulan dari item data yang saling berhubungan satu dengan yang lainnya yang diorganisasikan berdasarkan sebuah skema atau struktur tertentu, tersimpan di hardware komputer dan dengan software untuk melakukan manipulasi untuk kegunaan tertentu.

2. Rough set

Rough set merupakan salah satu teknik klasifikasi yang digunakan pada proses data mining, yang sumber datanya berasal dari sebuah database.

3. Knowledge

*Knowledge* (pengetahuan) merupakan sebuah informasi yang dihasilkan proses prediksi ynag kita lakukan. Knowledge yang dihasilkan dengan penerapan teori rough set tersebut berupa rule yang dijabarkan dalam bentuk IF....THEN.

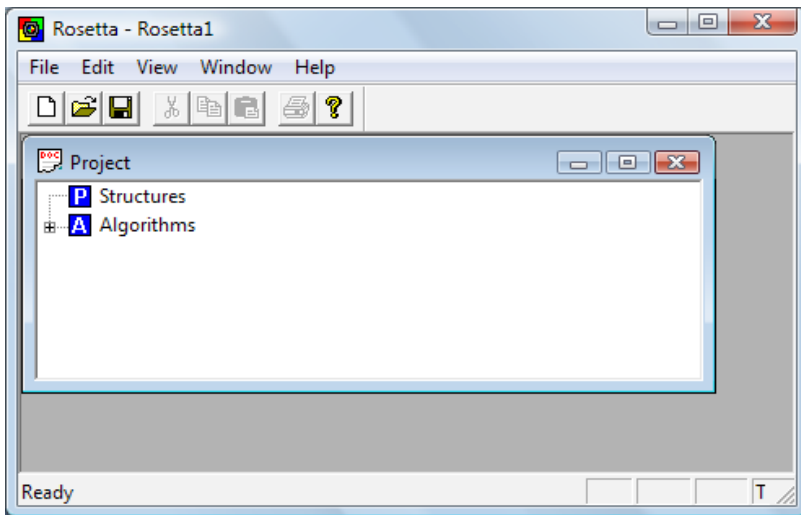
2. Pengujian Teory Rough Set

Untuk menguji kebenaran dari hasil pengolahan data yang dikerjakan secara manual kita dapat menggunakan salah satu *software* aplikasi *rough set* yaitu *Rosetta* dengan langkah-langkah sebagai berikut :

1. Seluruh variabel-variabel (terdiri dari attribut kondisi dan atribut keputusan) yang digunakan untuk memprediksi tingkat kelulusan siswa dalam ujian nasional disimpan pada microsoft excel dengan nama file DATA NILAI.xlsx seperti yang terlihat pada Gambar 2.
2. Buka aplikasi Rosetta, kemudian klik menu File → New, sehingga muncul tampilan seperti Gambar 3.

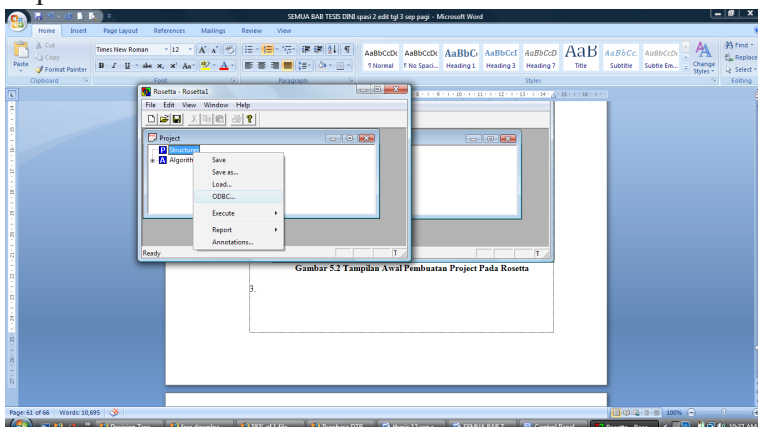
A	B	C	D	E	F	G
EXAMPLE	RATA2 NILAI RAPOR	RATA2 NILAI PRA UN	RATA2 NILAI UN	KET	tgkt kelulusan	
1						
2	S1	C	C	1	100%	
3	S2	C	D	1	75%	
4	S3	C	B	1	100%	
5	S4	C	D	1	75%	
6	S5	C	B	0	50%	
7	S6	C	D	1	75%	
8	S7	C	C	1	100%	
9	S8	C	C	1	100%	
10	S9	B	D	1	75%	
11	S10	B	B	1	100%	
12	S11	C	D	1	75%	
13	S12	C	D	1	75%	
14	S13	B	D	0	25%	
15	S14	B	D	1	75%	
16	S15	A	C	1	100%	
17	S16	C	C	1	100%	
18	S17	B	C	1	100%	
19	S18	C	D	1	75%	
20	S19	C	D	1	75%	
21	S20	C	D	1	75%	
22	S21	C	D	1	75%	

Gambar 2 File Data Nilai pada Ms. Excel



Gambar 3 Tampilan New Project Pada Rosetta

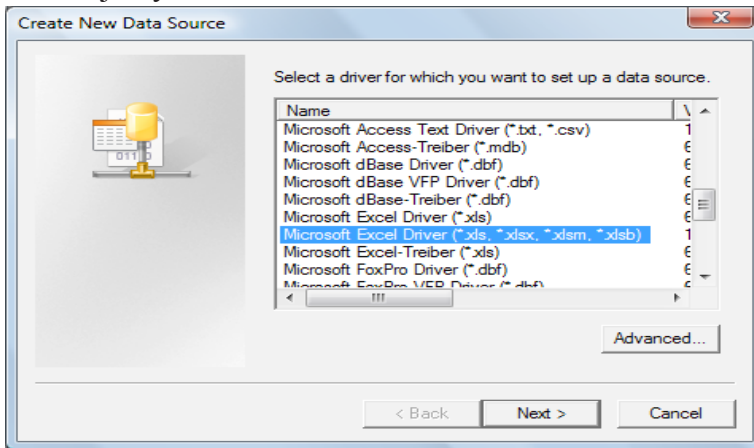
3. Selanjutnya insert file UJI B pada langkah 1 ke Project pada Rosetta, dengan langkah sebagai berikut :
  - a. Klik kanan pada *Structure*, pilih ODBC, seperti yang terlihat pada Gambar 3.



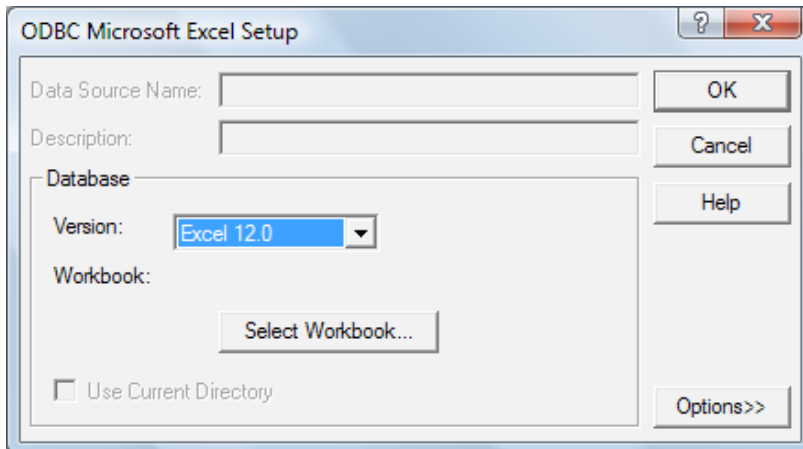
Gambar 3 Import File pada Rosetta

- b. Klik Tombol Open Database, lalu klik tombol *New*, kemudian pilih sumber file yang akan dipakai yaitu pilih *Microsoft Excel Driver* seperti yang terlihat pada Gambar 4. , setelah dipilih klik *Next*,

- c. Klik Browse untuk mencari file UJI B yang akan di insert ke Rosetta, kemudian Klik Next, selanjutnya *Finish*
- d. Kemudian muncul jendela ODBC Ms Excel *Setup* seperti yang terlihat pada Gambar 5, kemudian klik tombol *Select Workbook* dan cari file sumber pada langkah 1 (Gambar 6), selanjutnya klik OK.

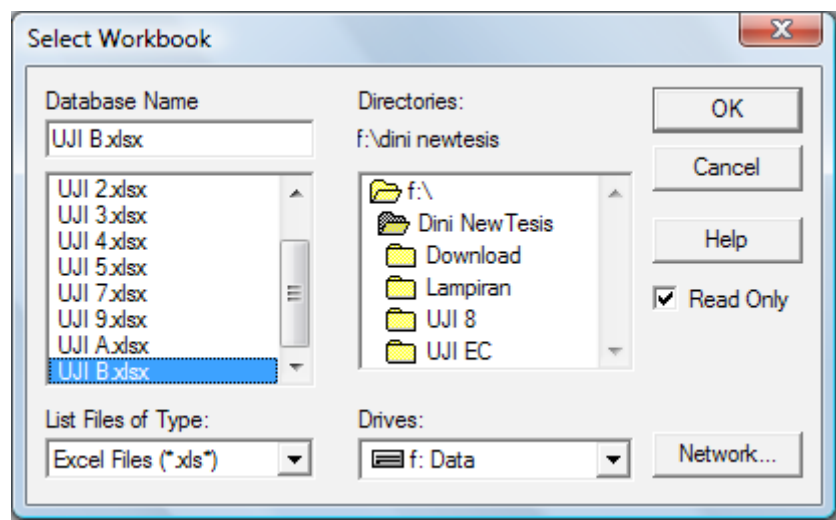


Gambar 4 Pilihan Data source yang tersedia pada Rosetta



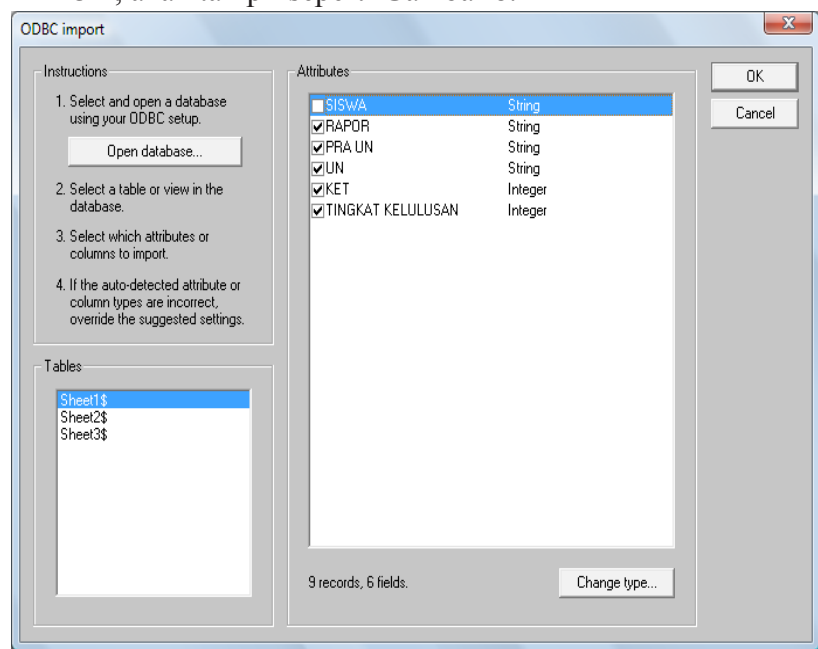
Gambar 5 Jendela ODBC Ms Excel Setup





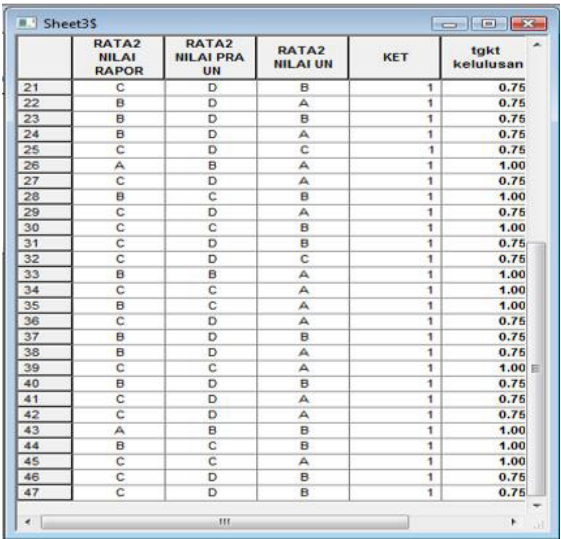
Gambar 6 Jendela Select Workbook

- e. Pada jendela ODBC Import pilih atribut-atribut yang akan digunakan seperti terlihat pada Gambar 7, kemudian klik OK, akan tampil seperti Gambar 8.



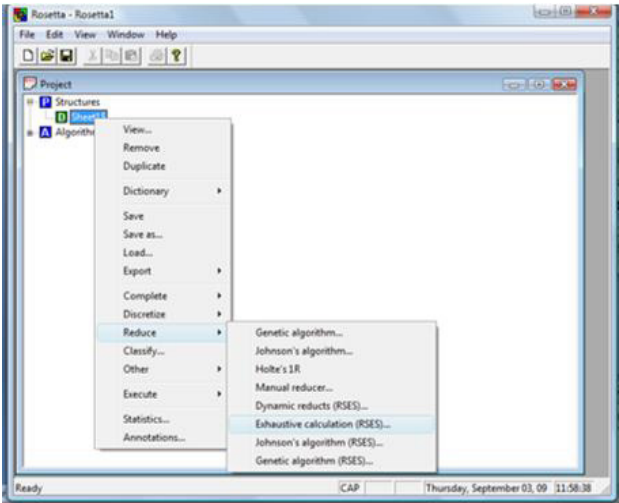
Gambar 7 Jendela ODBC import

- 4. Lakukan proses *Reduct* seperti gambar 9 dan hasil proses Reduct seperti yang terlihat pada Gambar 10
- 5. Lakukan *Generating Rules* seperti gambar 11 dan hasil prosesnya seperti yang terlihat pada Gambar 12. Pada langkah ini, terlihat bahwa hasil proses *Generating rule* berupa *knowledge* atau *rule* dalam bentuk IF...THEN.

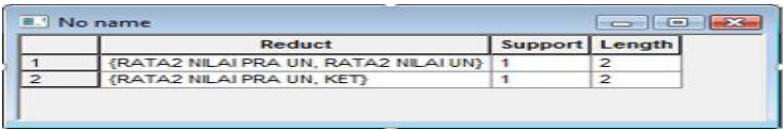


	RATA2 NILAI RAPOR	RATA2 NILAI PRA UN	RATA2 NILAI UN	KET	tgkt kelulusan
21	C	D	B	1	0.75
22	B	D	A	1	0.75
23	B	D	B	1	0.75
24	B	D	A	1	0.75
25	C	D	C	1	0.75
26	A	B	A	1	1.00
27	C	D	A	1	0.75
28	B	C	B	1	1.00
29	C	D	A	1	0.75
30	C	C	B	1	1.00
31	C	D	B	1	0.75
32	C	D	C	1	0.75
33	B	B	A	1	1.00
34	C	C	A	1	1.00
35	B	C	A	1	1.00
36	C	D	A	1	0.75
37	B	D	B	1	0.75
38	B	D	A	1	0.75
39	C	C	A	1	1.00
40	B	D	B	1	0.75
41	C	D	A	1	0.75
42	C	D	A	1	0.75
43	A	B	B	1	1.00
44	B	C	B	1	1.00
45	C	C	A	1	1.00
46	C	D	B	1	0.75
47	C	D	B	1	0.75

Gambar 8 Hasil Import File Excel pada Rosetta



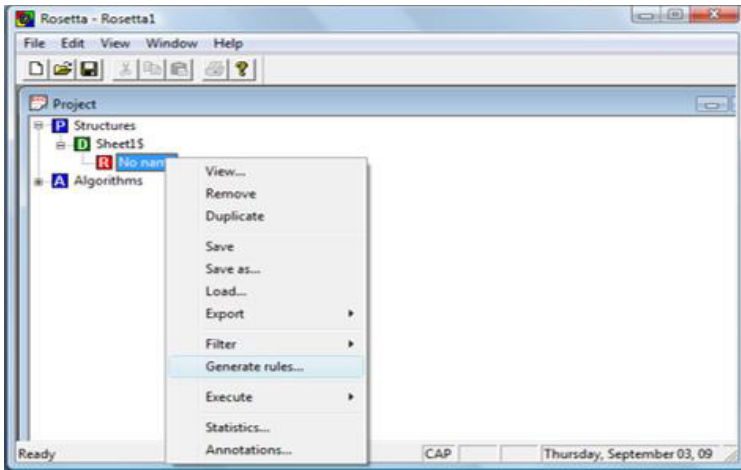
Gambar 9 Proses Reduct pada Rosetta



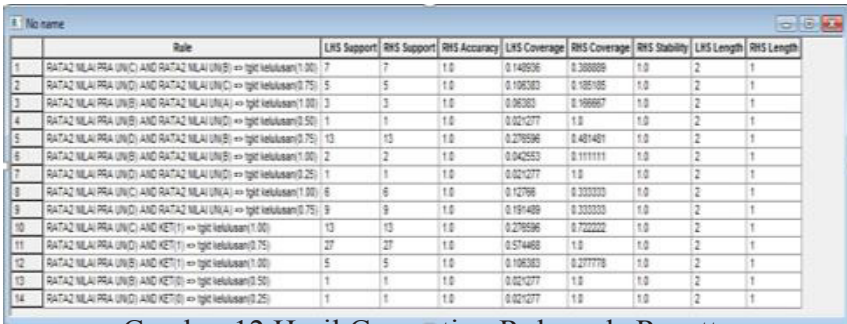
The screenshot shows a window titled 'No name' with a table containing two columns: 'Reduct' and 'Support'. The 'Reduct' column lists two sets of attributes: {RATA2 NILAI PRA UN, RATA2 NILAI UN} and {RATA2 NILAI PRA UN, KET}. The 'Support' column lists the value 1 for both. A 'Length' column on the right shows the value 2 for both.

	Reduct	Support	Length
1	{RATA2 NILAI PRA UN, RATA2 NILAI UN}	1	2
2	{RATA2 NILAI PRA UN, KET}	1	2

Gambar 10 Hasil Reduct pada Rosetta



Gambar 11 Proses Generating Rule pada Rosetta



The screenshot shows a window titled 'No name' with a table containing 14 rows of data. The columns are: 'Rule', 'LRS Support', 'RHS Support', 'RHS Accuracy', 'LRS Coverage', 'RHS Coverage', 'RHS Stability', 'LRS Length', and 'RHS Length'. The 'Rule' column lists 14 rules, each with a unique identifier and a set of attributes. The other columns contain numerical values representing various metrics for each rule.

Rule	LRS Support	RHS Support	RHS Accuracy	LRS Coverage	RHS Coverage	RHS Stability	LRS Length	RHS Length
1 RATA2 NILAI PRA UN(C) AND RATA2 NILAI UN(B) => tgr kelulusan(1.00)	7	7	1.0	0.148506	0.388889	1.0	2	1
2 RATA2 NILAI PRA UN(D) AND RATA2 NILAI UN(C) => tgr kelulusan(0.75)	5	5	1.0	0.106383	0.185185	1.0	2	1
3 RATA2 NILAI PRA UN(B) AND RATA2 NILAI UN(A) => tgr kelulusan(1.00)	3	3	1.0	0.06383	0.188889	1.0	2	1
4 RATA2 NILAI PRA UN(B) AND RATA2 NILAI UN(D) => tgr kelulusan(0.50)	1	1	1.0	0.021277	1.0	1.0	2	1
5 RATA2 NILAI PRA UN(D) AND RATA2 NILAI UN(B) => tgr kelulusan(0.75)	13	13	1.0	0.278596	0.481481	1.0	2	1
6 RATA2 NILAI PRA UN(B) AND RATA2 NILAI UN(B) => tgr kelulusan(1.00)	2	2	1.0	0.042553	0.111111	1.0	2	1
7 RATA2 NILAI PRA UN(D) AND RATA2 NILAI UN(D) => tgr kelulusan(0.25)	1	1	1.0	0.021277	1.0	1.0	2	1
8 RATA2 NILAI PRA UN(C) AND RATA2 NILAI UN(A) => tgr kelulusan(1.00)	6	6	1.0	0.12796	0.333333	1.0	2	1
9 RATA2 NILAI PRA UN(D) AND RATA2 NILAI UN(A) => tgr kelulusan(0.75)	9	9	1.0	0.191489	0.333333	1.0	2	1
10 RATA2 NILAI PRA UN(C) AND KET(1) => tgr kelulusan(1.00)	13	13	1.0	0.278596	0.722222	1.0	2	1
11 RATA2 NILAI PRA UN(D) AND KET(1) => tgr kelulusan(0.75)	27	27	1.0	0.574488	1.0	1.0	2	1
12 RATA2 NILAI PRA UN(B) AND KET(1) => tgr kelulusan(1.00)	5	5	1.0	0.106383	0.277778	1.0	2	1
13 RATA2 NILAI PRA UN(B) AND KET(0) => tgr kelulusan(0.50)	1	1	1.0	0.021277	1.0	1.0	2	1
14 RATA2 NILAI PRA UN(D) AND KET(0) => tgr kelulusan(0.25)	1	1	1.0	0.021277	1.0	1.0	2	1

Gambar 12 Hasil Generating Rule pada Rosetta

Setelah dilakukan pengujian pada software Rosetta V1.4.41 untuk memprediksi tingkat kelulusan siswa dalam ujian nasional seperti yang sudah dijelaskan di atas ternyata hasil rule yang dihasilkan pada Rosetta (Gambar 12) hampir sama dengan rule yang dihasilkan pada proses manual penemuan knowledge menggunakan teory rough set.

### 3. Pembahasan Hasil Pengujian

Dari dua hasil pengujian yang telah dilakukan yaitu proses secara manual dan menggunakan software *Rosetta* dapat kita ambil sebuah kesimpulan bahwa hasil pengujian sangat baik karena *rule* yang dihasilkan hampir sama.

Berdasarkan *rule-rule* yang dihasilkan (lihat pada lampiran) secara umum dapat kita lihat bahwa pada umumnya, siswa yang memiliki riwayat nilai yang bagus pada semua atau pada mata pelajaran tertentu yang di Ujian Nasional akan memiliki kecendrungan untuk mendapatkan nilai yang bagus juga dalam Ujian Nasional, begitu juga sebaliknya.

Seperti yang kita ketahui bersama, pada umumnya hasil ujian nasional ini sering bertentangan dengan kenyataan lapangan. Sebagaimana diketahui bahwa realitas pendidikan (sekolah) di Tanah Air sangat beragam, apakah itu sarana-prasarana pendidikan, sumber daya guru, dan school leadership.

Perbedaan kualitas pendidikan yang begitu lebar sebagai akibat dari keterbatasan kemampuan pengelola pendidikan pada tingkat pusat, daerah, dan sekolah semakin menguatkan tuduhan masyarakat selama ini bahwa penggunaan instrumen UN untuk menentukan kelulusan (sertifikasi) dan seleksi berpotensi misleading, bias, dan melanggar keadilan dalam tes.

Perbedaan kualitas pendidikan yang begitu lebar sebagai akibat dari keterbatasan kemampuan pengelola pendidikan pada tingkat pusat, daerah, dan sekolah semakin menguatkan tuduhan masyarakat selama ini bahwa penggunaan instrumen UN untuk menentukan kelulusan (sertifikasi) dan seleksi berpotensi misleading, bias, dan melanggar keadilan dalam tes.

Tapi hasil pengujian ini belum dapat dikatakan *valid* karena pada penelitian ini hanya menggunakan salah satu teknik untuk memprediksi yaitu teori *rough set*. Oleh sebab itu perlunya dilakukan penelitian lanjutan untuk membandingkan hasil prediksi dan menentukan teori mana yang menghasilkan prediksi yang baik.

Keteguhan sikap Menteri Pendidikan Nasional (Mendiknas) untuk tetap mempertahankan praktik UN pada sistem pendidikan menengah patut dihormati. Namun, pandangan dan pemikiran kritis terhadap praktik ujian akhir itu harus diutarakan agar sasaran yang dibuat dapat lebih proporsional, terarah, dan pencapaiannya dapat

dimaksimalkan. Meskipun praktik ujian akhir dapat digunakan untuk mempengaruhi kualitas pendidikan, namun pada umumnya sering bertentangan dengan kenyataan lapangan.

Berdasarkan hasil penelitian yang telah penulis lakukan yaitu di SMA Negeri 5 Kota Banda Aceh, Provinsi Nanggroe Aceh Darussalam untuk memprediksi tingkat kelulusan siswa dalam Ujian Nasional dapat disimpulkan bahwa :

1. Terdapat dua faktor penting yang menentukan keberhasilan penggunaan dari *data mining* adalah *gudang data* yang berukuran besar dan terintegrasi dengan baik, serta pemahaman atau identifikasi yang baik terhadap proses *data mining*.
2. Untuk mendapatkan hasil prediksi yang baik, dibutuhkan data atau pengetahuan yang baik pula.
3. Pemilihan variabel (attribut kondisi dan attribut keputusan) yang akan digunakan dalam memprediksi juga sangat mempengaruhi *rule* atau *knowledge* yang dihasilkan.
4. Selain itu, pemilihan metode prediksi yang digunakan tentu juga sangat berpengaruh pada hasil prediksi ini. Pada penelitian ini, penulis hanya mencoba salah satu teknik yang digunakan dalam memprediksi yaitu teori *rough set*.
5. Hasil proses *generating rule* pada teori *rough set* yang sudah dilakukan berupa *rule* dalam bentuk IF... THEN.

## DAFTAR PUSTAKA

- [1] **Thearling and Kurt (2005)** , “*An Introduction To Data Mining*”, **Whitepaper**, <http://www.shore.net/~kht/dmwhite/dmwhite.htm>
- [2] **Sarjon, D., and Mohd, N., (2002)**, “*Mining Association Rules Using Rough Set and Association Rules Methods*” , **Proceedings of ICAIET 2002** , **Page 435-440** , **Kota Kinabalu , Sabah , June 17-18.**
- [3] **Shao, J., (1996)**, “*Using Rough Set For Rough Classification*”
- [4] **Syamsi, I , ”Pengambilan Keputusan dan Sistem Informasi”** , **Bumi Aksara , Jakarta , 1995**
- [5] **Two Crows Corporation, Introduction to Data Mining and Knowledge Discovery, Third Edition. 10500 Falss Road. Potomac. MD 20854 ( U.S.A )**
- [6] **Yanyi, Y., T.C Chaim., (2000)**, “*Rule Discovery Based on Rough Set Theory*” , **ISIF**